

THE JEWETT-KRIEGER CONSTRUCTION FOR TILINGS

JEAN BELLISSARD AND IAN PALMER

ABSTRACT. Given a random distribution of impurities on a periodic crystal, an equivalent uniquely ergodic tiling space is built, made of aperiodic, repetitive tilings with finite local complexity, and with configurational entropy close to the entropy of the impurity distribution. The construction is the tiling analog of the Jewett-Kreger theorem.

*In memory of Pierre Duclos,
a long time friend,
with refreshing views on facts of life,
who dedicated himself
to the art of research in science*

1. INTRODUCTION

In the early seventies, Jewett [10], Krieger [14] (see also Hansel-Raoult [8]) proved the following theorem

Theorem 1 (Jewett-Krieger). *Let (X, \mathcal{A}, μ, S) be an ergodic dynamical system. Then there is a compact space Y and a homeomorphism T on Y , which is uniquely ergodic and minimal, such that if ν denotes the unique T -invariant probability measure on Y and \mathcal{B} the Borel σ -algebra of Y , the ergodic dynamical system (Y, \mathcal{B}, ν, T) is equivalent to (X, \mathcal{A}, μ, S) . In addition, (Y, \mathcal{B}, ν, T) can be constructed in a such a way that its topological entropy is as close as wished from the Kolmogorov-Sinai entropy of (X, \mathcal{A}, μ, S) .*

In this statement, \mathcal{A} denotes a σ -algebra of subsets of X , μ denotes a probability measure on \mathcal{A} and $S : X \mapsto X$ is a μ -preserving invertible bi-measurable transformation. Ergodicity means that if $A \in \mathcal{A}$ is S -invariant, then $\mu(A)$ is either 0 or 1. Two such spaces are called *equivalent* if there are sets $A \in \mathcal{A}$ and $B \in \mathcal{B}$, such that $\mu(A) = \nu(B) = 1$, $S(A) = A$, $T(B) = B$ and a transformation $\phi : A \mapsto B$ that is invertible and bi-measurable such that $T \circ \phi = \phi \circ S$.

The present paper is devoted to constructing a tiling that is aperiodic, repetitive, uniquely ergodic and has finite local complexity (FLC), starting from a random array of impurities. It will be shown that the two tiling spaces are equivalent and can be constructed so that the configurational entropy of the tiling space is arbitrarily close to the entropy of the original

Work supported in part by NSF Grants DMS-0600956, DMS-0901514 and CRC 701 of the DFG (German Research Council).

distribution of impurities. This construction is the analog for tiling spaces of the construction made by Krieger in his proof of Theorem 1. A more precise description of the result will be given in Section 2 and is summarized in Theorem 2.

In Theorem 1, two properties must be emphasized. First, the theorem gives a more systematic construction of uniquely ergodic dynamical systems with positive entropy. A classical article by Oxtoby [16] gives a review on unique ergodicity. It is known, for instance, that substitution sequences are uniquely ergodic [17]. A similar situation occurs for substitution tilings in higher dimensions. Hahn and Katznelson [7] showed that uniquely ergodic systems with arbitrarily large entropy exist. Recently, uniquely ergodic minimal tilings with positive configurational entropy were constructed by Cortez [4] using an explicit machinery making the method more practical.

The second important property is that it is possible to identify any given ergodic dynamical system, modulo sets of zero probability, with a uniquely ergodic minimal system having almost the same entropy. In the case of tilings this latter situation has not yet been investigated. This “prevalence” of unique ergodicity may have interesting consequences in a famous challenging problem of Mathematical Physics, the Anderson transition (see, for instance, [2] for a short review of this problem). Theorem 2 allows the Anderson problem to be reduced to a problem on a deterministic tiling.

2. NOTATIONS AND MAIN RESULT

2.1. Impurities. Impurities usually arise in a periodic crystal. In some cases, impurities are essential in accounting for the electronic properties of the crystal, such as the case of semiconductors like silicon or gallium-arsenide. Without impurities, this crystal is perfectly periodic; i.e. the atomic nuclei are located on a discrete subset \mathcal{L}_0 that is invariant by a closed, discrete subgroup of \mathbb{R}^d isomorphic to \mathbb{Z}^d . Therefore, in the perfect crystal the atomic arrangement is completely prescribed. Let \mathfrak{A} be the list of possible atomic impurities that can be found in such a crystal, to which a special letter denoted by 0 will be added. This list, which is finite, will be called an *alphabet*. A distribution of impurities can be seen as a sequence $\xi = (\xi_{\mathbf{x}})_{\mathbf{x} \in \mathcal{L}_0} \in \mathfrak{A}^{\mathcal{L}_0}$ in which $\xi_{\mathbf{x}}$ represents the atomic species that is located at \mathbf{x} , so that $\xi_{\mathbf{x}} = 0$ means that the atomic nucleus located at \mathbf{x} is the one that should normally be found at \mathbf{x} in the perfect crystal (absence of impurity at \mathbf{x}). A distribution of impurities is normally submitted to the laws of Thermodynamics. However, if their concentration is very small, such as in semiconductors used for electronics, the laws of Thermodynamics imply that the distribution of impurities is Poissonian with a very good accuracy. Thus, there is a probability measure μ on \mathcal{A} such that the $\xi_{\mathbf{x}}$ ’s can be seen as identical independent random variables with common distribution μ . This process is defined on the probability space $X = (\mathfrak{A}^{\mathcal{L}_0}, \mathcal{A}, \mathbb{P}_{\mu})$ where \mathcal{A} is the σ -algebra generated by cylinder sets and \mathbb{P}_{μ} is the probability induced by μ . The group \mathbb{Z}^d acts in a natural way on X through the shift; i.e. if $\mathbf{y} \in \mathbb{Z}^d$ then $(\tau^{\mathbf{y}}\xi)_{\mathbf{x}} = \xi_{\mathbf{x}-\mathbf{y}}$. A well known elementary result of ergodic theory shows that this action is measure preserving and ergodic. In particular, given any $R > 0$, with probability one there is a cube of size R in the crystal in which no impurity can be found.

In the following, the *entropy* of μ is the Shannon entropy, defined by

$$S(\mu) = - \sum_{a \in \mathfrak{A}} \mu(a) \ln \mu(a)$$

which, in this case, coincides with the Kolmogorov-Sinai entropy of the dynamical system.

2.2. Tilings. In this work only a very restricted class of tessellations, called *c-tilings*, will be considered: tilings of \mathbb{R}^d with *decorated* tiles in the shape of unit hypercubes (called its *geometric support*) centered at a point in \mathbb{Z}^d . Examples of a more general approach to tilings can be found in [12, 1, 15, 19]. A tile will be decorated (i) by a letter in the alphabet \mathfrak{A} , called its *color* and (ii) by its collar, namely the set of (colored) tiles touching it. A tile decorated only by its color will be called *colored* while it will be called *collared* if the collar is added to the decoration. A *c-tiling* (or *tessellation*) is a collection of colored tiles such that the interiors of their geometric supports have pairwise empty intersections while the union of their geometric supports is \mathbb{R}^d . An example of a tiling is given by a distribution of impurities on \mathbb{Z}^d . A *patch* is a union of a finite number of distinct colored tiles contained in the tiling. The geometric support of the patch is the union of geometric supports of the tiles contained in the patch. In this work, only patches whose geometric support is a hypercube with center in \mathbb{Z}^d will be considered. An equivalence class of (colored, collared) tiles (*resp.* patches) modulo translations is called a (colored, collared) proto-tile (*resp.* proto-patch).

Given a c-tiling \mathcal{T} , let \mathcal{P}_n be the set of proto-patches of size $2n + 1$. Obviously this set is finite, since the alphabet is finite. Hence a c-tiling has *finite local complexity* (FLC) or has *finite type* [12, 1, 15, 19]. Let $\mathcal{P} = \mathcal{P}(\mathcal{T})$ be the (disjoint) union of the \mathcal{P}_n 's and call it the *dictionary* of \mathcal{T} or its *atlas*. The *tiling space* of \mathcal{T} , also called the *transversal*, is the set of all tilings in \mathbb{R}^d having the same dictionary. For any pair of integers $m < n$, there is a natural restriction map $\pi_{n,m} : \mathcal{P}_n \mapsto \mathcal{P}_m$. It follows immediately that the tiling space is homeomorphic to $\Xi = \varprojlim (\mathcal{P}_n, \pi_{n,m})$. In particular Ξ is a compact totally disconnected space [12].

The group \mathbb{Z}^d acts in an obvious way on Ξ since two tilings that only differ by a translation have the same dictionary. Moreover, it is elementary to show that \mathbb{Z}^d acts by homeomorphisms. Thus (Ξ, \mathbb{Z}^d) is a topological dynamical system. A tiling \mathcal{T} is called *aperiodic* whenever $\mathcal{T} + \mathbf{x} = \mathcal{T}$ for some $\mathbf{x} \in \mathbb{Z}^d$ if and only if $\mathbf{x} = 0$. In this case, Ξ has no isolated points and is thus a Cantor set.

Proposition 1. *Let \mathcal{T} be a c-tiling and Ξ its tiling space. Then there is a canonical map $\phi : \Xi \mapsto \mathfrak{A}^{\mathbb{Z}^d}$ which is continuous, one-to-one, and conjugates the \mathbb{Z}^d -actions.*

Proof. Since Ξ is given by the previous inverse limit, an element $T \in \Xi$ is an infinite sequence $(p_n)_{n \in \mathbb{N}}$ with $p_n \in \mathcal{P}_n$ and $\pi_{n+1,n}(p_{n+1}) = p_n$. In particular, p_n is a hypercubic patch of size $(2n + 1)$ occupying the center of p_{n+1} . Thus the union $\bigcup_n p_n$ defines a tiling, namely an infinite family of colored tiles, thus a point $\phi_n(T) \in \mathfrak{A}^{\mathbb{Z}^d}$. From this construction, it follows immediately that ϕ is one-to-one. Moreover it is also continuous. For indeed, a basis for the topology of $\mathfrak{A}^{\mathbb{Z}^d}$ is provided by a cylinder set based on some hypercubic patches. The inverse image of such a patch is an open set of the basis defining the topology of the inverse limit. The construction shows immediately why ϕ conjugates the \mathbb{Z}^d -actions. \square

A tiling \mathcal{T} is called *repetitive* if, given any patch p , there is an $R > 0$ such that in any hypercube of size R there is a copy (modulo translation) of p . Clearly, a distribution of impurities of the kind described in Section 2.1 is almost surely NOT repetitive. Thanks to a well known theorem [18], a tiling is repetitive if and only if its tiling space is *minimal* with respect to the \mathbb{Z}^d -action.

Given a tiling \mathcal{T} with dictionary \mathcal{P} , let $|\mathcal{P}_n|$ be the cardinality of \mathcal{P}_n . Then the *configurational entropy* is given by

$$H_c(\mathcal{T}) = \limsup_{n \rightarrow \infty} \frac{\ln |\mathcal{P}_n|}{(2n+1)^d}$$

2.3. Main Result. The main result of this paper is given by

Theorem 2. *Let $X = (\mathfrak{A}^{\mathcal{L}_0}, \mathcal{A}, \mathbb{P}_\mu, \mathbb{Z}^d)$ be the ergodic system provided by a random distribution of impurities on the periodic set \mathcal{L}_0 , where μ is a probability measure on the alphabet \mathfrak{A} . Then, for any $\epsilon > 0$ there exist a tiling \mathcal{T} which is aperiodic, repetitive and FLC, such that*

- (i) *its tiling space (Ξ, \mathbb{Z}^d) is uniquely ergodic*
- (ii) *its configurational entropy is within ϵ from the entropy $S(\mu)$ of the measure μ*
- (iii) *the ergodic system $Y = (\Xi, \mathcal{B}, \nu, \mathbb{Z}^d)$ is equivalent to X , where \mathcal{B} is the Borel σ -algebra of Ξ and ν is the unique \mathbb{Z}^d -invariant measure.*

Our construction is essentially the proof given by Krieger [14]. The method used in the proof consists in defining inductively a family of patches of increasing size, generating the dictionary of the tiling that is to be constructed by keeping only the patches with empirical distribution close to the distribution provided by μ in the system of impurities X . By controlling how close this is, it will be possible to prove that there is a unique invariant ergodic measure on the tiling space.

Proof of the equivalence: This invariant measure will be shown to be close to \mathbb{P}_μ (see Lemma 4.4 and the Remark in Section 5.1). As a consequence of Proposition 1, $\mathbb{P}_\mu \circ \phi(\Xi) > 0$. Using the translation invariance of $\phi(\Xi)$ and the ergodicity of \mathbb{P}_μ it will follow that $\mathbb{P}_\mu \circ \phi(\Xi) = 1$. By the unique ergodicity, the restriction of \mathbb{P}_μ to $\phi(\Xi)$ is the image of the unique invariant measure on Ξ . The equivalence of the two spaces X, Y follows. \square

Acknowledgements. This work was funded by NSF Grants DMS-0600956 and DMS-0901514. Both authors thank the CRC 701 in Bielefeld and the DFG (German Research Council) for financial support during the completion of this work. I.P. wishes to thank A. Grigo for several helpful conversations. Both authors are indebted to P. Arnoux, J. Kellendonk and L. Sadun for comments and useful new references.

3. ESTIMATES FROM PROBABILITY AND INFORMATION THEORY

An excellent reference for this section is the book by Cover and Thomas [5], especially Chapter 11. Given a finite set \mathfrak{A} (the *alphabet*) of cardinality $|\mathfrak{A}|$, let $\mathcal{M}_1(\mathfrak{A})$ denote the space of probabilities on \mathfrak{A} . Fix an integer $l > 1$ (the *word length*) and let \mathfrak{A}^l denote the set of words w of length l with letters $w_i \in \mathfrak{A}$ for $i = 1, \dots, l$. For any $\mu \in \mathcal{M}_1(\mathfrak{A})$, let μ^l denote the corresponding product measure on \mathfrak{A}^l .

Definition 3.1 (Empirical Distribution). For $w \in \mathfrak{A}^l$, the *empirical distribution* $\pi^w \in \mathcal{M}_1(\mathfrak{A})$ of w gives the occurrence frequencies of the letters in \mathfrak{A} :

$$\pi^w(a) := \frac{1}{l} \sum_{i=1}^l \delta_{aw_i}$$

where δ_{aw_i} is the point measure equal to 1 when $a = w_i$ and 0 otherwise.

Definition 3.2 (Total Variation Distance). The (*total variation*) *distance* between two measures $\nu, \nu' \in \mathcal{M}_1(\mathfrak{A})$ is

$$d(\nu, \nu') := \sum_{a \in \mathfrak{A}} |\nu(a) - \nu'(a)|$$

There are three related entropies to be considered for measures ν, ν' on \mathfrak{A} :

$$\begin{array}{ll} \text{(Shannon) Entropy} & S(\nu) := - \sum_{a \in \mathfrak{A}} \nu(a) \ln(\nu(a)) \\ \text{Relative Entropy} & D_{KL}(\nu, \nu') := \sum_{a \in \mathfrak{A}} \nu(a) \text{Log}\left(\frac{\nu(a)}{\nu'(a)}\right) \\ \text{Cross Entropy} & S(\nu, \nu') := - \sum_{a \in \mathfrak{A}} \nu(a) \ln(\nu'(a)) \end{array}$$

Each of the three quantities is nonnegative; $S(\nu)$ and $S(\nu, \nu')$ are so because $\nu(a) \leq 1$ implies $\ln(\nu(a)) \leq 0$, and $D_{KL}(\nu, \nu')$ is by a straightforward convexity argument [5, Theorem 2.6.3]. It follows from the definitions that $S(\nu, \nu') - S(\nu) = D_{KL}(\nu, \nu')$, and thus

$$S(\nu, \nu') \geq S(\nu) \tag{*}$$

The following elementary lemma provides an estimate on $S(\nu, \nu') - S(\nu')$ when the measures ν, ν' are not too far apart.

Lemma 3.3. *For any measures $\nu, \nu' \in \mathcal{M}_1(\mathfrak{A})$ and $\epsilon \leq \min \{\nu'(a) \mid a \in \mathfrak{A}\}$, if $d(\nu, \nu') < \epsilon$ then*

$$|S(\nu, \nu') - S(\nu')| < \epsilon \ln\left(\frac{1}{\epsilon}\right)$$

Proof.

$$\begin{aligned} |S(\nu, \nu') - S(\nu')| &= \left| - \sum_{a \in \mathfrak{A}} (\nu(a) - \nu'(a)) \ln(\nu'(a)) \right| \\ &\leq \left(\sum_{a \in \mathfrak{A}} |\nu(a) - \nu'(a)| \right) \max_{a \in \mathfrak{A}} \left\{ \ln\left(\frac{1}{\nu'(a)}\right) \right\} \\ &\leq \epsilon \ln\left(\frac{1}{\epsilon}\right) \end{aligned} \quad \square$$

The property of having the same empirical distribution is an equivalence relation on the words in \mathfrak{A}^l . For $w \in \mathfrak{A}^l$, let

$$\chi_w := \left\{ w' \in \mathfrak{A}^l \mid \pi^w(a) = \pi^{w'}(a) \quad \forall a \in \mathfrak{A} \right\}$$

be the equivalence class of w under this relation. For a fixed alphabet \mathfrak{A} and word length l the number of distinct equivalence classes is bounded by a polynomial in the word length l (of degree $|\mathfrak{A}|$) [5, Theorem 11.1.1], and the probability of a word w only depends on its class:

Lemma 3.4 ([5, Theorem 11.1.2]). *Given a word $w \in \mathfrak{A}^l$ and a measure $\mu \in \mathcal{M}_1(\mathfrak{A})$,*

$$\mu^l(w) = e^{-lS(\pi^w, \mu)}$$

Thus the cardinality of a class χ_w is given by:

$$|\chi_w| = \mu^l(\chi_w) e^{lS(\pi^w, \mu)} \quad (3.1)$$

With (3.1) and Lemma 3.3 it is now possible to give an estimate on the number of words whose empirical distributions are ϵ -close to a given probability $\mu \in \mathcal{M}_1(\mathfrak{A})$.

Lemma 3.5. *Given a fixed alphabet \mathfrak{A} and word length l , for any measure $\mu \in \mathcal{M}_1(\mathfrak{A})$ and $\epsilon \leq \min \{ \mu(a) \mid a \in \mathfrak{A} \}$, let*

$$W(\mu, \epsilon) := \left\{ w \in \mathfrak{A}^l \mid d(\pi^w, \mu) < \epsilon \right\}$$

be the set of words of length l whose empirical distributions are ϵ -close to μ . Then

$$\mu^l(W(\mu, \epsilon)) e^{l[S(\mu) - \epsilon \ln(\frac{1}{\epsilon})]} \leq |W(\mu, \epsilon)| \leq \mu^l(W(\mu, \epsilon)) e^{l[S(\mu) + \epsilon \ln(\frac{1}{\epsilon})]}$$

Proof. Let $\mathcal{N}(\mu, \epsilon)$ be the set of distinct classes χ_w for which w satisfies $d(\pi^w, \mu) < \epsilon$. Then

$$\left| \ln(|W(\mu, \epsilon)|) - lS(\mu) \right| = \left| \ln \left(\sum_{\chi_w \in \mathcal{N}(\mu, \epsilon)} \mu^l(\chi_w) e^{l[S(\pi^w, \mu) - S(\mu)]} \right) \right|$$

by Lemma 3.4, and by Lemma 3.3

$$\mu^l(W(\mu, \epsilon)) e^{-l\epsilon \ln(\frac{1}{\epsilon})} \leq \sum_{\chi_w \in \mathcal{N}(\mu, \epsilon)} \mu^l(\chi_w) e^{l[S(\pi^w, \mu) - S(\mu)]} \leq \mu^l(W(\mu, \epsilon)) e^{l\epsilon \ln(\frac{1}{\epsilon})}$$

so the result follows. \square

In addition to the estimate on $|W(\mu, \epsilon)|$ above, it is possible to give a lower bound on $\mu^l(W(\mu, \epsilon))$. The following concentration inequality is a direct consequence of the Chernoff-Hoeffding Theorem [9].

Lemma 3.6. *For \mathfrak{A} , l , and μ as in the previous lemma and any $\epsilon > 0$,*

$$\mu^l(W(\mu, \epsilon)) > 1 - 2|\mathfrak{A}| e^{-2l\left(\frac{\epsilon}{|\mathfrak{A}|}\right)^2}$$

Proof. Because the expectation with respect to μ of each δ_{aw_i} is $\mu(a)$, for any fixed $a \in \mathfrak{A}$ the Chernoff-Hoeffding Theorem gives the bound

$$\mu^l \left(\left\{ w \in \mathfrak{A}^l \mid |\pi^w(a) - \mu(a)| \geq \epsilon \right\} \right) < 2e^{-2l\epsilon^2}$$

The result follows, since

$$\left\{ w \in \mathfrak{A}^l \mid d(\pi^w, \mu) \geq \epsilon \right\} \subseteq \bigcup_{a \in \mathfrak{A}} \left\{ w \in \mathfrak{A}^l \mid |\pi^w(a) - \mu(a)| \geq \frac{\epsilon}{|\mathfrak{A}|} \right\}$$

□

Let $W^C(\mu, \epsilon)$ be the complement of $W(\mu, \epsilon)$ in \mathfrak{A}^l , and for ease of notation, let $\gamma := 2l \left(\frac{\epsilon}{|\mathfrak{A}|} \right)^2$. It is now possible to give estimates for measures on \mathfrak{A}^l .

Lemma 3.7. *For \mathfrak{A} , l , μ and ϵ as in the previous lemma, let $\hat{\mu}$ be the normalized restriction of μ^l to $W(\mu, \epsilon) \subset \mathfrak{A}^l$. Then*

$$d(\mu^l, \hat{\mu}) < 4|\mathfrak{A}|e^{-\gamma} \quad \text{and} \quad 0 < lS(\mu) - S(\hat{\mu}) < 4|\mathfrak{A}|le^{-\gamma} \ln \left(\frac{1}{\epsilon} \right)$$

Proof. The first inequality is a consequence of Lemma 3.6:

$$\begin{aligned} d(\mu^l, \hat{\mu}) &= \sum_{w \in W(\mu, \epsilon)} \left| \mu^l(w) - \hat{\mu}(w) \right| + \sum_{w \in W^C(\mu, \epsilon)} \mu^l(w) \\ &= 2 \left(1 - \mu^l(W(\mu, \epsilon)) \right) < 4|\mathfrak{A}|e^{-\gamma} \end{aligned}$$

For the second statement, $\hat{\mu}$ is the conditional probability of $\mu^l(w)$ subject to the condition $d(\pi^w, \mu) < \epsilon$. The difference $S(\mu^l) - S(\hat{\mu})$ must be positive, since the unconditioned probability μ^l must have the greater entropy; since $S(\mu^l) = lS(\mu)$, the lower bound is established and it remains to give the upper bound. Combining the technique of Lemma 3.3 with the previous inequality gives

$$\begin{aligned} \left| S(\mu^l, \hat{\mu}) - S(\hat{\mu}) \right| &\leq \sum_{w \in W(\mu, \epsilon)} \left| \mu^l(w) - \hat{\mu}(w) \right| \ln \left(\frac{1}{\hat{\mu}(w)} \right) \\ &\leq 4|\mathfrak{A}|e^{-\gamma} \max_{w \in W(\mu, \epsilon)} \ln \left(\frac{\mu^l(W(\mu, \epsilon))}{\mu^l(w)} \right) \end{aligned}$$

Since $\mu^l(w) \leq 1$ for all words w , it is enough to find a lower bound for $\mu^l(w)$, and since $\mu^l(w) = \prod_{i=1}^l \mu(w_i) \geq \epsilon^l$, it follows that $|S(\mu^l, \hat{\mu}) - S(\hat{\mu})| \leq 4|\mathfrak{A}|e^{-\gamma} \ln \left(\frac{1}{\epsilon^l} \right)$. By inequality (*) above, $S(\mu^l, \hat{\mu}) \geq S(\mu^l)$, so

$$S(\mu^l) - S(\hat{\mu}) \leq S(\mu^l, \hat{\mu}) - S(\hat{\mu}) \leq 4|\mathfrak{A}|le^{-\gamma} \ln \left(\frac{1}{\epsilon} \right)$$

□

4. CONSTRUCTION OF ANDERSON-PUTNAM COMPLEX

4.1. Alphabets and Measures. In this section, a sequence of alphabets $\{\mathfrak{A}_n\}_{n=0}^\infty$ is constructed inductively so that for each $n \in \mathbb{N}$, the alphabet \mathfrak{A}_n is a family of words of length l_n with letters in the alphabet \mathfrak{A}_{n-1} . The word lengths l_n will satisfy certain constraints to ensure that the system $\{\mathfrak{A}_n\}_{n=0}^\infty$ has the desired properties.

Let \mathfrak{A}_0 be an initial finite alphabet with $|\mathfrak{A}_0| \geq 2$, and fix a measure $\mu_0 \in \mathcal{M}_1(\mathfrak{A}_0)$ that is strictly positive on all $a \in \mathfrak{A}_0$. Fix a parameter $\beta \in (0, \frac{1}{2})$. For each $n = 0, \dots, \infty$, assume

- \mathfrak{A}_n is a finite alphabet,
- $\mu_n \in \mathcal{M}_1(\mathfrak{A}_n)$ is strictly positive for all $a \in \mathfrak{A}_n$,

then,

- choose an integer

$$l_{n+1} > \max_{a \in \mathfrak{A}_n} \left\{ (\mu_n(a))^{-\frac{1}{\beta}} \right\} \quad (\text{Constraint 1})$$

- and let $\epsilon_n = l_{n+1}^{-\beta}$ (so that $\epsilon_n < \min \{\mu_n(a) \mid a \in \mathfrak{A}_n\}$).

Then construct the next generation by

- (1) letting

$$\mathfrak{A}_{n+1} := \left\{ w \in \mathfrak{A}_n^{l_{n+1}} \mid d(\pi^w, \mu_n) < \epsilon_n \right\} \quad (4.1)$$

- (2) and letting

$$\mu_{n+1} := \frac{1}{\mu_n^{l_{n+1}}(\mathfrak{A}_n)} \mu^{l_{n+1}}|_{\mathfrak{A}_{n+1}}$$

be the normalized product measure $\mu_n^{l_{n+1}}$ after restricting to $\mathfrak{A}_{n+1} \subset \mathfrak{A}_n^{l_{n+1}}$.

Let $L_n := l_n \cdots l_1$ be the number of letters of \mathfrak{A}_0 in a word of \mathfrak{A}_n , let $l_0 = L_0 = 1$, and, for ease of notation, let $\gamma_n := 2l_{n+1} \left(\frac{\epsilon_n}{|\mathfrak{A}_n|} \right)^2$. Finally, define a projection $\phi_n : \mathcal{M}_1(\mathfrak{A}_n^{l_{n+1}}) \rightarrow \mathcal{M}_1(\mathfrak{A}_n)$ by

$$\nu \mapsto \sum_{w \in \mathfrak{A}_n^{l_{n+1}}} \nu(w) \pi^w$$

and for any integer $m > n$, let $\Phi_n^m := \phi_n \circ \cdots \circ \phi_{m-1}$.

Since $\epsilon_n < \min \{\mu_n(a) \mid a \in \mathfrak{A}_n\}$, each letter of \mathfrak{A}_n must occur at least once in each word of \mathfrak{A}_{n+1} (because $|\pi^w(a) - \mu_n(a)| < \epsilon_n < \mu_n(a)$ for each $a \in \mathfrak{A}_n$), and, as a further consequence, $l_{n+1} > |\mathfrak{A}_n|$.

Lemma 4.1. *The family $\{\pi^w \mid w \in \mathfrak{A}^l\}$ of empirical distributions over the words in \mathfrak{A}^l is a lattice in $\mathcal{M}_1(\mathfrak{A})$, and for neighboring points π^w and $\pi^{w'}$, $d(\pi^w, \pi^{w'}) = \frac{2}{l}$.*

Proof. Obvious. □

Proposition 4.2. *For each $n \in \mathbb{N}$, the set \mathfrak{A}_n is nonempty.*

Proof. Given the collection of $|\mathfrak{A}|$ point measures $\{\delta_a\}_{a \in \mathfrak{A}}$, the measure with the greatest distance from the δ_a is the uniform distribution $\frac{1}{|\mathfrak{A}|} \sum_{a \in \mathfrak{A}} \delta_a$; this maximum distance is

$$d\left(\delta_{a'}, \frac{1}{|\mathfrak{A}|} \sum_{a \in \mathfrak{A}} \delta_a\right) = \sum_{b \in \mathfrak{A}} \left| \delta_{a'b} - \frac{1}{|\mathfrak{A}|} \sum_{a \in \mathfrak{A}} \delta_{ab} \right| = 2 \left(1 - \frac{1}{|\mathfrak{A}|}\right)$$

By the previous lemma, the distance to the center of one of the cells is therefore $\frac{4}{l} \left(1 - \frac{1}{|\mathfrak{A}|}\right)$. With

$$\epsilon = l^{-\beta} > \frac{4}{l} \left(1 - \frac{1}{|\mathfrak{A}|}\right) \quad (4.2)$$

there is guaranteed to be a word $w \in \mathfrak{A}^l$ satisfying $d(\pi^w, \mu) < \epsilon$ for any $\mu \in \mathcal{M}_1(\mathfrak{A})$ (for instance, any $l > 4$ satisfies this inequality regardless of $|\mathfrak{A}|$). \square

Lemma 4.3. *Recall that $L_n = l_n \cdots l_1$ and $S(\mu^l) = lS(\mu)$. Then, for each $n \in \mathbb{N}$,*

$$\left| \frac{1}{L_n} \ln(|\mathfrak{A}_n|) - S(\mu_0) \right| \leq \frac{\epsilon_{n-1}}{L_{n-1}} \ln\left(\frac{1}{\epsilon_{n-1}}\right) + \sum_{k=1}^{n-1} \frac{4|\mathfrak{A}_{k-1}| e^{-\gamma_{k-1}}}{L_{k-1}} \ln\left(\frac{1}{\epsilon_{k-1}}\right)$$

Proof.

$$\left| \frac{1}{L_n} \ln(|\mathfrak{A}_n|) - S(\mu_0) \right| \leq \frac{1}{L_{n-1}} \left| \frac{1}{l_n} \ln(|\mathfrak{A}_n|) - S(\mu_{n-1}) \right| + \sum_{k=1}^{n-1} \frac{1}{L_k} |S(\mu_k) - l_k S(\mu_{k-1})|$$

and, since $\mu_{n-1}^l(\mathfrak{A}_n) \leq 1$ for each n , the result follows by Lemmas 3.5 and 3.7. \square

Unique Ergodicity. An essential ingredient of the present construction is the contraction property of the projections Φ_n^{n+1} (see [6, 3]), demonstrated in the following lemma. The contractions of the Φ_n^{n+1} will cause the space of probability measures to collapse to a point in the inverse limit.

Lemma 4.4. *Given arbitrary choices of $\nu_m \in \mathcal{M}_1(\mathfrak{A}_m)$ for each m , the sequence of measures $\{\Phi_n^m \nu_m\}_{m=n+1}^\infty$ in $\mathcal{M}_1(\mathfrak{A}_n)$ is Cauchy for any n . Thus there is a unique $\tilde{\mu}_n$ in each $\mathcal{M}_1(\mathfrak{A}_n)$ such that*

$$\lim_{m \rightarrow \infty} \Phi_n^m \nu_m = \tilde{\mu}_n$$

and, furthermore,

$$\phi_n(\tilde{\mu}_{n+1}) = \tilde{\mu}_n$$

Proof. For any $\nu, \nu' \in \mathcal{M}_1(\mathfrak{A}_{n+1})$,

$$\begin{aligned} d(\Phi_n^{n+1} \nu, \Phi_n^{n+1} \nu') &= \sum_{a \in \mathfrak{A}_n} \left| \sum_{w \in \mathfrak{A}_{n+1}} [\pi^w(a) - \mu(a)] [\nu(w) - \nu'(w)] \right| \\ &\leq \sum_{w \in \mathfrak{A}_{n+1}} d(\pi^w, \mu) |\nu(w) - \nu'(w)| \leq \epsilon_n d(\nu, \nu') \end{aligned}$$

By induction, given $m' > m > n$,

$$d\left(\Phi_n^m \nu_m, \Phi_n^{m'} \nu_{m'}\right) \leq \prod_{k=n}^{m-1} \epsilon_k d\left(\nu_m, \Phi_m^{m'} \nu_{m'}\right) \leq 2 \prod_{k=n}^{m-1} \epsilon_k$$

Since $\epsilon_k = l_{k+1}^{-\beta}$ decays exponentially, $\{\Phi_n^m \nu_m\}_{m>n}$ is Cauchy, converging to a measure $\tilde{\mu}_n$.

Furthermore,

$$\begin{aligned} d\left(\Phi_n^{n+1} \tilde{\mu}_{n+1}, \tilde{\mu}_n\right) &\leq d\left(\Phi_n^{n+1} \tilde{\mu}_{n+1}, \Phi_n^m \nu_m\right) + d\left(\Phi_n^m \nu_m, \tilde{\mu}_n\right) \\ &\leq \epsilon_n d\left(\tilde{\mu}_{n+1}, \Phi_{n+1}^m \nu_m\right) + d\left(\Phi_n^m \nu_m, \tilde{\mu}_n\right) \end{aligned}$$

for all $m > n$ so $\Phi_n^{n+1} \tilde{\mu}_{n+1} = \tilde{\mu}_n$. The choice of the original sequence $\{\nu_m\}_{m=0}^\infty$ was arbitrary, so the sequence $\{\tilde{\mu}_n\}_{n=0}^\infty$ is unique. \square

4.2. Construction of Tilings. The system of words $\{\mathfrak{A}_n\}_{n=0}^\infty$ can be used to label a tiling of \mathbb{R}^d by d -cubic tiles. By fixing an ordering on d -cubic configurations of tiles, a one-to-one correspondence between words of a given length and configurations of a corresponding size is established. The constraints on the word lengths l_n are all lower bounds, so it is possible to impose the additional constraint that for each $n \in \mathbb{N}$,

$$l_n = (2r_n + 1)^d \quad \text{for some } r_n \in \mathbb{N} \quad (\text{Constraint 2})$$

In this way it will be possible to map each word in \mathfrak{A}_n to a d -cube of volume $L_n := l_n \cdots l_1$, the number of letters of the initial alphabet \mathfrak{A}_0 in a word in \mathfrak{A}_n .

For $n = 0, \dots, \infty$, an n -patch is the interior of a d -cube of volume $L_n := l_n \cdots l_1$; each n -patch is the interior of the union of the closures of l_n $(n-1)$ -patches, arranged $(2r_n + 1)$ on a side. A 0-patch is also called a *tile*. The location of an n -patch is determined by its center, assumed to be the origin unless stated otherwise, and always an element of the integer lattice \mathbb{Z}^d seen as a subset of \mathbb{R}^d . For an n -patch P , let $\mathcal{L}_P \subset \mathbb{Z}^d$ be the point set consisting of the centers of the l_n $(n-1)$ -patches that comprise P .

An *order* on \mathcal{L}_P is a bijection $\{1, \dots, l_n\} \rightarrow \mathcal{L}_P$. Fix such an order $\sigma_{n,0}$ on \mathcal{L}_P for each $n \in \mathbb{N}$, and for an n -patch P centered at \mathbf{x} , let $\sigma_{n,\mathbf{x}}(i) = \sigma_{n,0}(i) + \mathbf{x}$ so that the ordering is translation invariant in each generation n .

A *labeled*, or *colored*, n -patch is a pair (P, w) , where $w \in \mathfrak{A}_n$. If $w = w_1 \dots w_{l_n}$, the $(n-1)$ -patch at $\sigma_{n,0}(i)$ is labeled by w_i . By recursion, this assigns a label in \mathfrak{A}_0 to each of the L_n tiles in P . Let \mathcal{P}_n be the collection of all labeled n -patches, which is in one-to-one correspondence with \mathfrak{A}_n . For $n = 0, \dots, \infty$, since \mathfrak{A}_n is invariant under permutations of the letters in a word, \mathcal{P}_n is invariant under permutations of the $(n-1)$ -patches in an n -patch.

To ensure that a tiling formed from the system of patches $\{\mathcal{P}_n\}_{n=0}^\infty$ forces its border [11], it is sufficient to collar the tiles. The *geometric base* of an n -patch P is the interior of the union of the closures of the tiles in P , and the *geometric boundary* of a patch is the boundary of its geometric base. For any patch $P \in \mathcal{P}_n$, the geometric base is restricted to that of the $(2r_n - 1)^d$ $(n-1)$ -patches at the center of P ; the label w on P now also determines the $(n-1)$ -patches adjacent to P . Thus the labeled patch P now represents a collared patch—a patch whose nearest neighbors are known.

The *boundary* ∂P of an n -patch P can be decomposed into a disjoint union of codimension- k boundary components, each of which is labeled by the 2^k n -patches adjacent to it. For $\mathbf{x} \in \partial P$, $\mathbf{y} \in \partial P'$, let $\mathbf{x} \sim_n \mathbf{y}$ when \mathbf{x} and \mathbf{y} occupy the same position in a boundary component with the same 2^k labels.

Following the philosophy of the Anderson-Putnam complex as presented in [3], for $n \in \mathbb{N}$, the branched manifolds \mathcal{B}_n are quotients of the disjoint union of the collared, colored n -patches in \mathcal{P}_n by the relation \sim_n . Because \mathcal{P}_n is invariant under permutations of the $(n-1)$ -patches in an n -patch, each $(n-1)$ -patch shares each of its boundary components with every other $(n-1)$ patch in \mathcal{P}_{n-1} in some n -patch in \mathcal{P}_n .

The projections $\varphi_{n-1}^n : \mathcal{B}_n \rightarrow \mathcal{B}_{n-1}$ of the Anderson-Putnam complex are determined by the labeling of n -patches by $(n-1)$ -patches. The singular locus of \mathcal{B}_n is the set of points belonging to more than one n -patch; the *regions* of \mathcal{B}_n are the connected components of \mathcal{B}_n after removing the singular locus, one for each collared, colored n -patch. Each region is tiled by $(n-1)$ -patches, and the image under φ_{n-1}^n of a point \mathbf{x} in a region is determined by its location in the $(n-1)$ -patch that contains it. Points in the boundaries of multiple $(n-1)$ -patches are sent to the corresponding points in the singular locus of \mathcal{B}_{n-1} . The Hull Ω is the inverse limit $\varprojlim (\mathcal{B}_n, \varphi_{n-1}^n)$.

5. PROPERTIES OF THE HULL

The proof of the theorem is completed by establishing the properties of Ω . For any $\mathcal{T} \in \Omega$, because there are finitely many tiles of a given size and they always meet full-face to full-face, \mathcal{T} has FLC. For all $n \in \mathbb{N}$, every $(n-1)$ -patch occurs in every n -patch, so \mathcal{T} is also repetitive, implying minimality of the dynamical system (Ω, \mathbb{R}^d) . It follows that Ω is the closure of the orbit of a repetitive FLC tiling under the \mathbb{R}^d -action, so every tiling in Ω is aperiodic if any one is [13]. To construct an aperiodic tiling in Ω it is sufficient to choose a sequence of words $\{w^{(n)}\}_{n=0}^\infty$, with $w^{(n)} \in \mathfrak{A}_n$ so that $w_i^{(n)} = w_i^{(n-1)}$ for i such that $\sigma_{n,0}(i) = 0$, choosing each $w^{(n)}$ so that the n -patches do not allow translation symmetry.

5.1. Unique Ergodicity. In [3] it was shown that the top dimension singular homology group $H_d(\mathcal{B}_n, \mathbb{R})$ has a canonical orientation, and that the positive weights on the \mathcal{B}_n are the elements of the positive cone in $H_d(\mathcal{B}_n, \mathbb{R})$. Furthermore, with maps $(\varphi_n^{n+1})_* : H_d(\mathcal{B}_n, \mathbb{R}) \rightarrow H_d(\mathcal{B}_{n+1}, \mathbb{R})$ induced by the projections $\varphi_n^{n+1} : \mathcal{B}_{n+1} \rightarrow \mathcal{B}_n$, the invariant measures on the transversal Ξ of Ω are determined by the elements of the inverse limit $\varprojlim (H_d(\mathcal{B}_n, \mathbb{R}), (\varphi_n^{n+1})_*)$.

A positive weight on \mathcal{B}_n is a distribution on the regions of \mathcal{B}_n satisfying Kirchhoff's Law. This is equivalent to the condition that, given a face of codimension-1, the sum of the weights on the region on one side of the face be equal to the sum of the weights on the region on the opposite side. Due to the invariance of the \mathcal{P}_n under permutations of the $(n-1)$ -patches in an n -patch, for any measure in $\mathcal{M}_1(\mathfrak{A}_n) = \mathcal{M}_1(\mathcal{P}_n)$, either sum is equal to a conditional probability given by the product of the measure on the two regions.

Thus the weights of mass one on \mathcal{B}_n are given by the probabilities on \mathcal{P}_n , and the projections Φ_n^{n+1} defined on the \mathfrak{A}_n are the same as the induced maps $(\varphi_n^{n+1})_*$. By Lemma 4.4 there

is a unique sequence of measures $\tilde{\mu}_n$ on the \mathfrak{A}_n such that $\Phi_n^{n+1}\tilde{\mu}_{n+1} = \tilde{\mu}_n$ for $n = 0, \dots, \infty$, determining the unique invariant measure on Ξ .

Proposition 5.1. *The configurational entropy of $\mathcal{T} \in \Omega$ is positive and can be made arbitrarily close to $S(\mu_0)$.*

Proof. Let $h_n := \frac{1}{L_n} \ln(|\mathfrak{A}_n|)$. Since \mathfrak{A}_n is a proper subset of $\mathfrak{A}_{n-1}^{l_n}$ for each n , $\{h_n\}_{n \in \mathbb{N}}$ is a decreasing sequence of positive numbers. By Lemma 4.3, if $h_\infty := \lim_{n \rightarrow \infty} h_n$, then

$$|h_\infty - S(\mu_0)| \leq \sum_{n=0}^{\infty} \frac{4|\mathfrak{A}_n|e^{-\gamma_n}}{L_n} \ln\left(\frac{1}{\epsilon_n}\right)$$

so $|h_\infty - S(\mu_0)| < \sum_{n=0}^{\infty} \frac{1}{L_n}$. Since

$$S(\mu_0) = \ln\left(e^{S(\mu_0)}\right) = \sum_{n=1}^{\infty} \frac{1}{n} \left(1 - e^{-S(\mu_0)}\right)^n$$

$\sum_{n=0}^{\infty} \frac{1}{L_n} < S(\mu_0)$ provided $l_{n-1} > \frac{n}{(n-1)(1-e^{-S(\mu_0)})}$ for each $n \in \mathbb{N}$, which is trivial since l_n grows exponentially in n . Thus $h_\infty > 0$, and by choosing the sequence of word lengths l_n to be larger, the difference between $S(\mu_0)$ and the configurational entropy can be made arbitrarily small. \square

Remark. For $n = 0, \dots, \infty$, the distance between μ_n and the product measure $\mu_{n-1}^{l_n}$ is small and is controlled by the parameter β . By Lemma 3.7, $d(\tilde{\mu}_n, \mu_n) \leq \sum_{m>n} 4|\mathfrak{A}_m|e^{-\gamma_m}$, with $\gamma_m = \frac{2}{|\mathfrak{A}_m|^2} l_m^{1-2\beta}$; just as the parameter δ can be adjusted to control the growth rate of the words, β can be adjusted to control the deviation of $\tilde{\mu}_n$ from μ_n . Since Ω is uniquely ergodic, the metric entropy of $\tilde{\mu}$ is equal to the configurational entropy and hence the difference between these and the entropy of the initial choice of measure μ_0 can be made arbitrarily small.

REFERENCES

- [1] Anderson, J.E., Putnam, I.F.: Topological invariants for substitution tilings and their associated C^* -algebras. *Ergod. Th. & Dynam. Sys.*, **18** 1998, 509–537.
- [2] Bellissard, J.: Random Matrix Theory and the Anderson Model. *J. Stat. Phys.*, **116** (1-4), (2004), 739–754.
- [3] Bellissard, J., Benedetti, R., Gambaudo, J.-M.: Spaces of tilings, finite telescopic approximations and gap-labeling. *Communications in Mathematical Physics*, **261**(3) 2006, pp. 1–41.
- [4] Cortez, M.I.: \mathbb{Z}^d -Toeplitz arrays. *Discrete Contin. Dyn. Syst.* **15**(3) 2006, pp. 859–881.
- [5] Cover, T.M., Thomas, J.A.: *Elements of Information Theory*, 2nd ed., Wiley, 2006.
- [6] Gambaudo, J.-M., Martens, M.: Algebraic topology for minimal cantor sets. *Annales Henri Poincare*, **7**(3) 2006, pp. 423–446. (*A preprint was available in 1999.*)
- [7] Hahn, F., Katznelson, Y.: On the entropy of uniquely ergodic transformations. *Transactions of the American Mathematical Society*, **126**(2) Feb. 1967, pp. 335–360.
- [8] Hansel, G., Raoult, J.-P.: Ergodicity, uniformity and unique ergodicity. *Indiana Univ. Math. J.*, **23**, (1973/74), 221–237.
- [9] Hoeffding, W.: Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, **58**(301) Mar. 1963, pp. 13–30.
- [10] Jewett, R.I.: The prevalence of uniquely ergodic systems. *J. Math. Mech.*, **19** 1969/1970, pp. 717–729.
- [11] Kellendonk, J.: Non-commutative geometry of tilings and gap labelling. *Reviews in Mathematical Physics*, **7**(07) 1995, pp. 1133–1180.
- [12] Kellendonk, J.: Local structure of tilings and their integer group of coinvariants. *Commun. Math. Phys.*, **187**, (1997), 115–157.
- [13] Kellendonk, J., Putnam, I.F.: Tilings, C^* -algebras and K -theory. in *Directions in Mathematical Quasicrystals*, CRM Monograph Series, Volume **13**, M.P. Baake, R.V. Moody, Eds., AMS, Providence, RI: 2000, pp. 177–206.
- [14] Krieger, W.: On unique ergodicity. *Proc. Sixth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1970/1971)*, Vol. II: *Probability theory*, Univ. California Press, Berkeley, Calif., 1972, pp. 327–346.
- [15] Lagarias, J.C.: Geometric models for quasicrystals I. Delone sets of finite type. *Discrete Comput. Geom.*, **21** 1999, pp. 161–191.
- [16] Oxtoby, J. C.: Ergodic sets. *Bull. Amer. Math. Soc.*, **58** 1952, pp. 116–136.
- [17] Queffelec, M.: *Substitution Dynamical Systems - Spectral Analysis*. Lecture Notes in Mathematics, 1294. Springer 1987.
- [18] Radin, C., Wolff, M: Space tilings and local isomorphism. *Geom Dedicata*, **42**, (1992), 355–360.
- [19] Sadun, L.: *Topology of Tiling Spaces*. University Lecture Series, 46. American Mathematical Society, Providence, RI, 2008.

GEORGIA INSTITUTE OF TECHNOLOGY, SCHOOL OF MATHEMATICS, ATLANTA GA 30332-0160

E-mail address: jeanbel@math.gatech.edu

GEORGIA INSTITUTE OF TECHNOLOGY, SCHOOL OF MATHEMATICS, ATLANTA GA 30332-0160

E-mail address: icpalm@math.gatech.edu